Evaluating Analytical Impact of Observed Instabilities in Diffusion MRI Pipelines

Gregory Kiar

2020-2021 Canadian Open Neuroscience Platform Student Research Fellow,
2018-2021 NSERC Alexander Graham Bell Fellow,
2017 McGill University Healthy Brains for Healthy Lives Fellow,
Ph.D. Candidate (Biomedical Engineering) McGill University,
M.S.E. (Biomedical Engineering) Johns Hopkins University,
B.Eng. (Biomedical & Electrical Engineering) Carleton University.

Meeting Contents

Outline of thesis
Summary of original contributions
Summary of the work to date
Summary of work remaining to be done
List of publications submitted/planned
Schedule for graduation

Thesis Objective

This thesis aims to...

1. evaluate the numerical stability of typical workflows,
2. quantify the implications for the reliability of neuroscience studies,
3. develop methods to increase their robustness.

Thesis Outline (Manuscripts)

- Intro & Background

- Chapter/Paper 1
- Chapter/Paper 2
- Chapter/Paper 3
- Chapter/Paper 4

- Discussion

Chapter 1: A Serverless Tool for Platform Agnostic Computational Experiment Management

Chapter 2: Comparing Perturbation Models for Evaluating Stability of Neuroimaging Pipelines

Chapter 3: Numerical Instabilities in Analytical Pipelines Compromise the Reliability of Network Neuroscience

Chapter 4: Augmenting Connectomics Datasets Through Aggregation of Numerically Unstable Derivatives

Chapter 1: A Serverless Tool for Platform Agnostic Computational Experiment Management

Clowdr

Chapter 2: Comparing Perturbation Models for Evaluating Stability of Neuroimaging Pipelines

Identifying Instabilities

Chapter 3: Numerical Instabilities in Analytical Pipelines Compromise the Reliability of Network Neuroscience

Impact of Instabilities

Chapter 4: Augmenting Connectomics Datasets Through Aggregation of Numerically Unstable Derivatives

Aggregating Unstable Derivatives

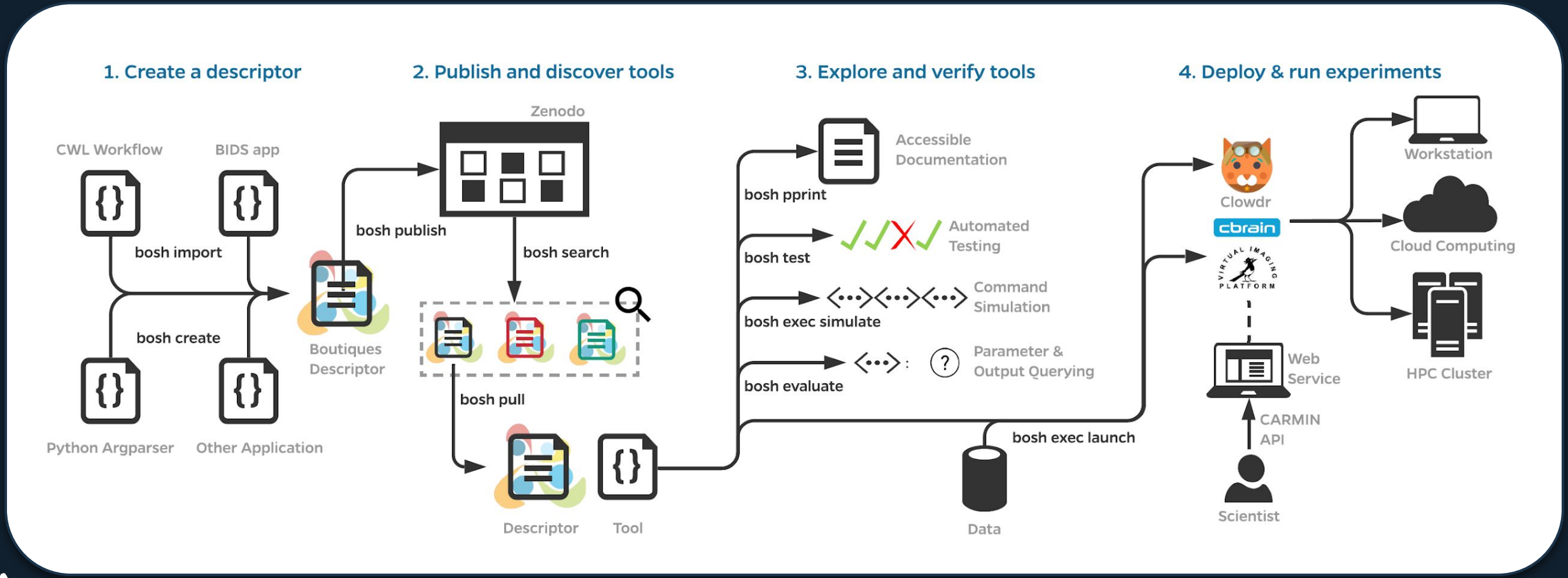
Chapter 1: A Serverless Tool for Platform Agnostic Computational Experiment Management

Complete (Published in Frontiers in Neuroinformatics)

Paper: https://doi.org/10.3389/fninf.2019.00012

Tool: https://doi.org/10.5281/zenodo.2537168

Clowdr Facilitates FAIR Tool Deployment



(Kiar, 2019)

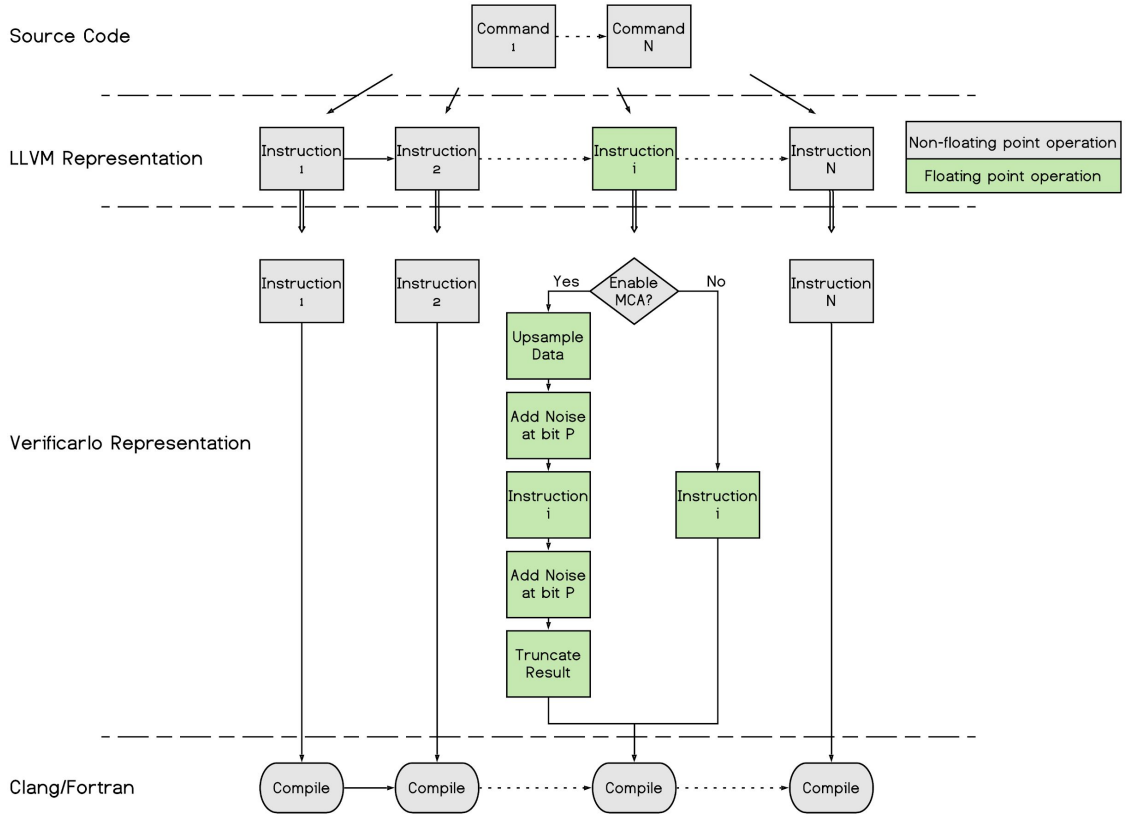
Chapter 2: Comparing Perturbation Models for Evaluating Stability of Neuroimaging Pipelines

Complete (published in IJHPCA)

Paper: https://doi.org/10.1177/1094342020926237

Dataset: https://doi.org/10.5281/zenodo.4033308

Inducing Instabilities with MCA



Base

✓	# 68.1	AMD64	Verificarlo

Environments (level 1)

✓	# 68.2	AMD64	Blas & Lapack
✓	# 68.3	AMD64	Python
✓	# 68.4	AMD64	Libmath

Environments (level 2)

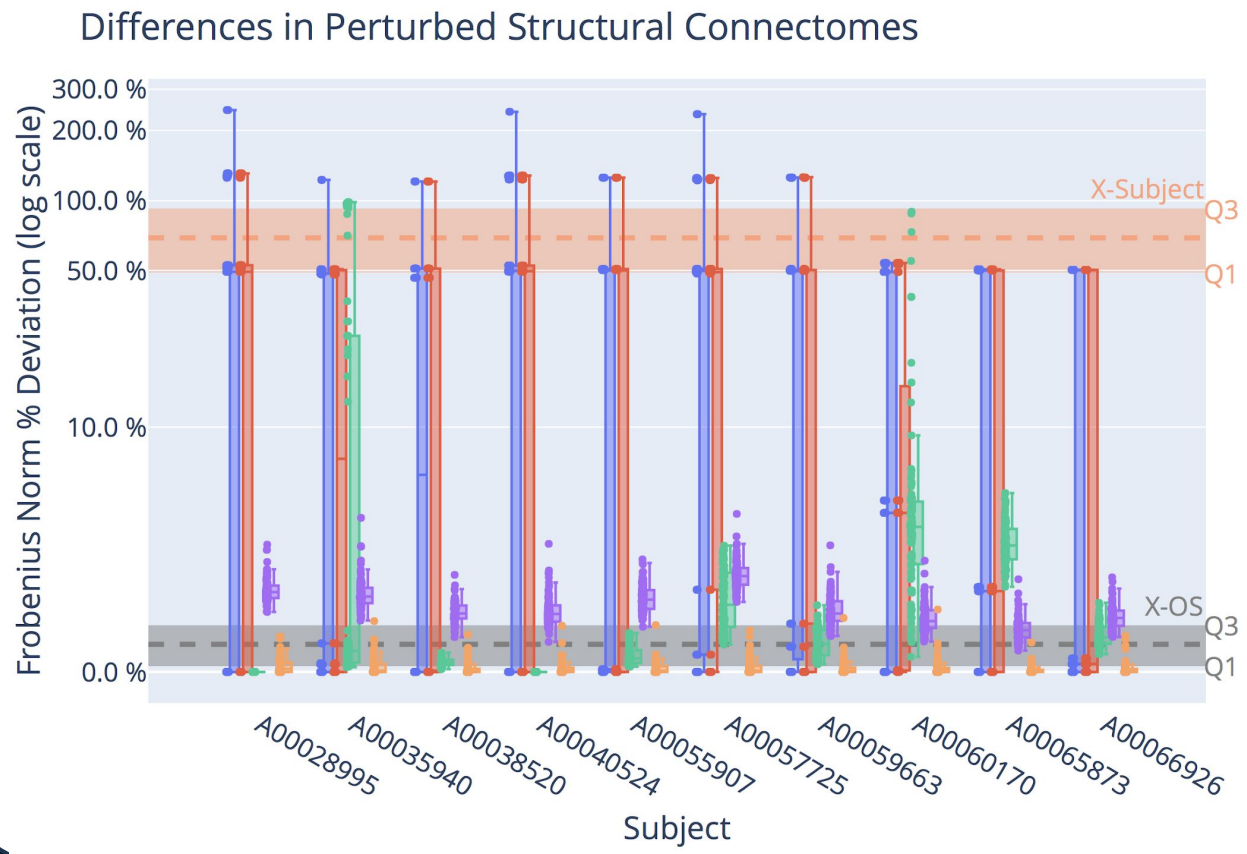
✓	# 68.5	AMD64	Python & Numpy

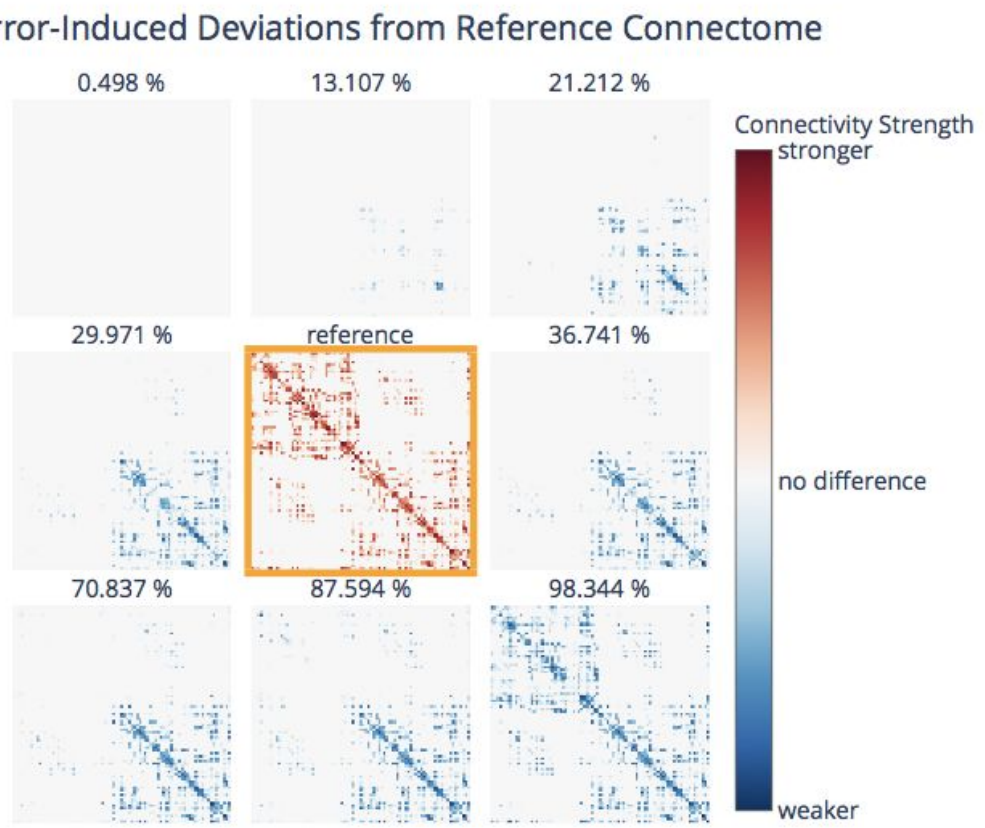
Applications

✓	# 68.6	AMD64	FSL
✓	# 68.7	AMD64	AFNI
✓	# 68.8	AMD64	AFQ (Dipy)

https://github.com/gkiar/fuzzy

Connectomes are (Variably) Variable





(Kiar, 2020)

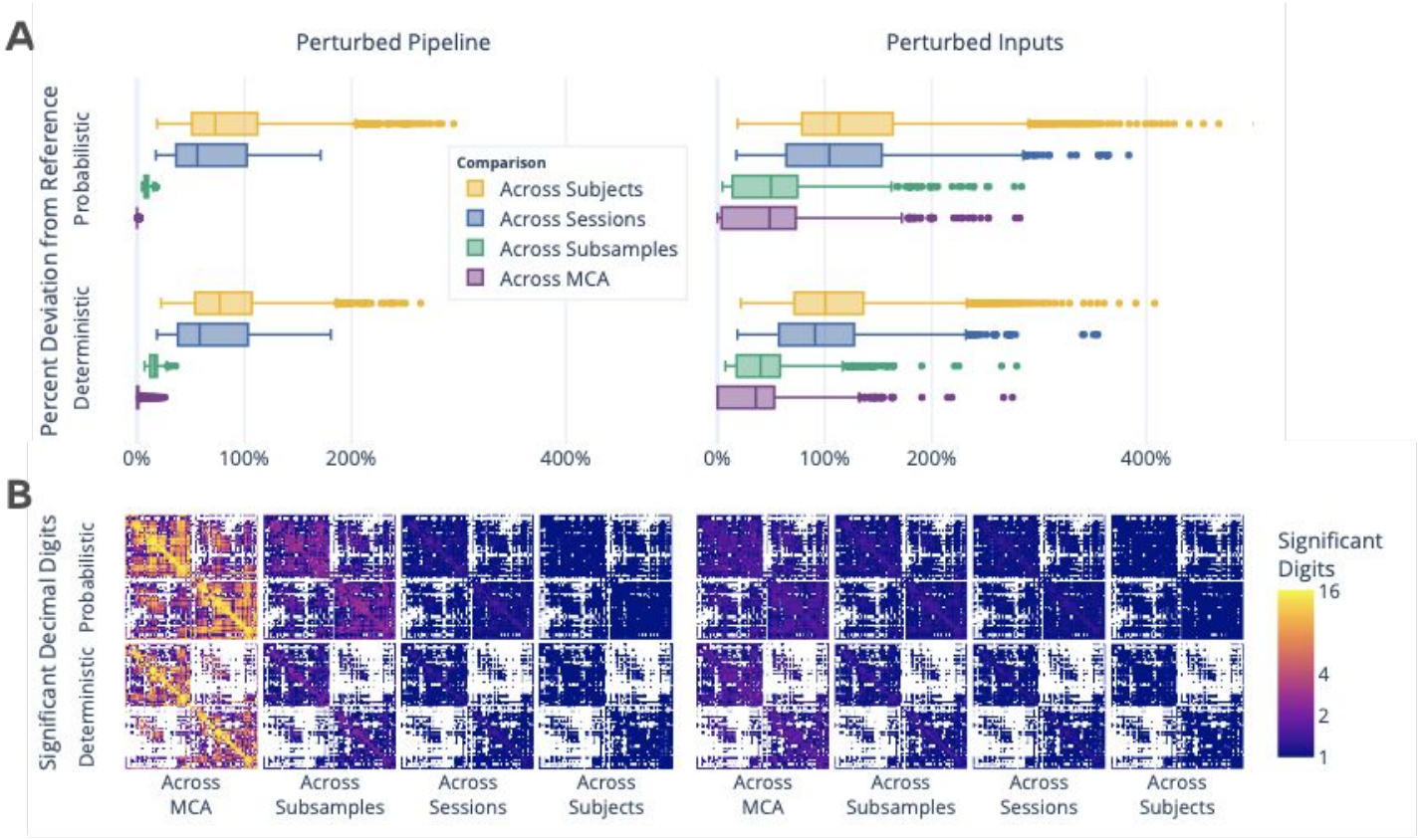
Chapter 3: Numerical Instabilities in Analytical Pipelines Compromise the Reliability of Network Neuroscience

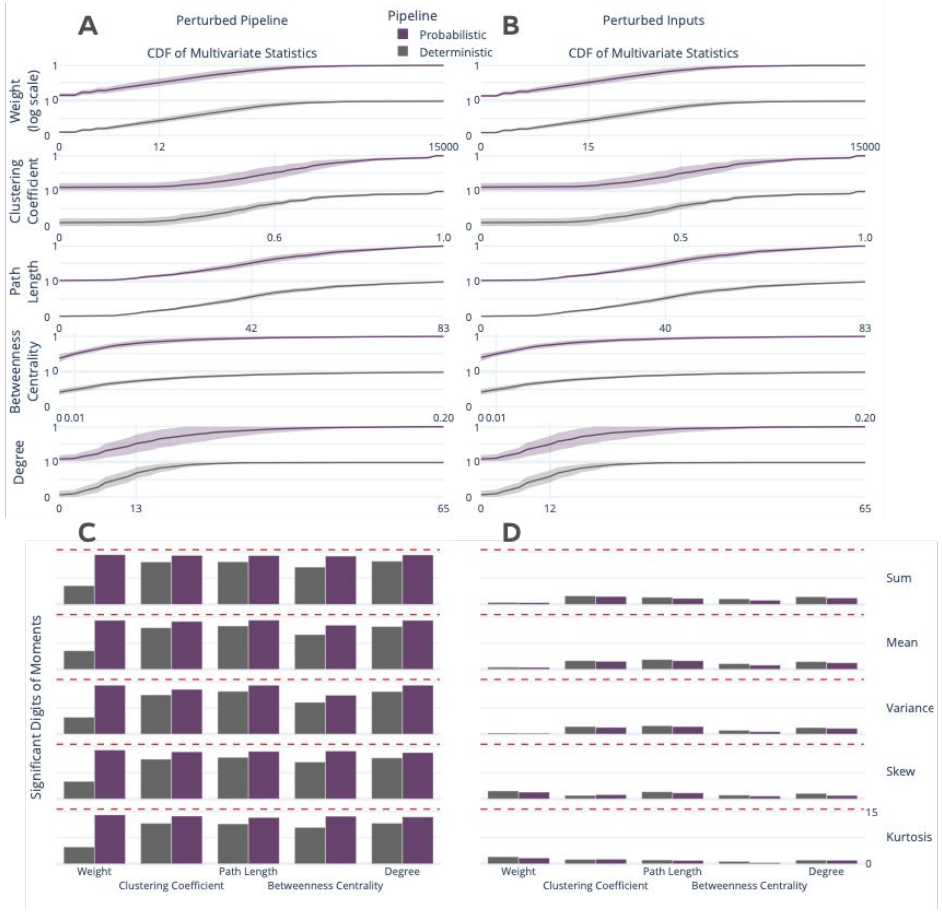
Work Complete (to be submitted to Nature)

Preprint: https://github.com/gkpapers/2020ImpactOfInstability

Dataset: https://doi.org/10.5281/zenodo.4041549

Instabilities Can Obscure Meaningful Signal





(Kiar, 2020)

Instabilities Can ALSO Increase Reliability

Such as increasing similarity across subjects (H_1) …

Comparison	Chance	Target	Reference Execution		Perturbed Pipeline		Perturbed Inputs	
			Det.	Prob.	Det.	Prob.	Det.	Prob.
H_1: Across Subjects	0.04	1.0	0.64	0.65	0.82	0.82	0.77	0.75
H_2: Across Sessions	0.5	0.5	1.00	1.00	1.00	1.00	0.88	0.85
H_3: Across Subsamples	0.5	0.5			0.99	1.00	0.71	0.61

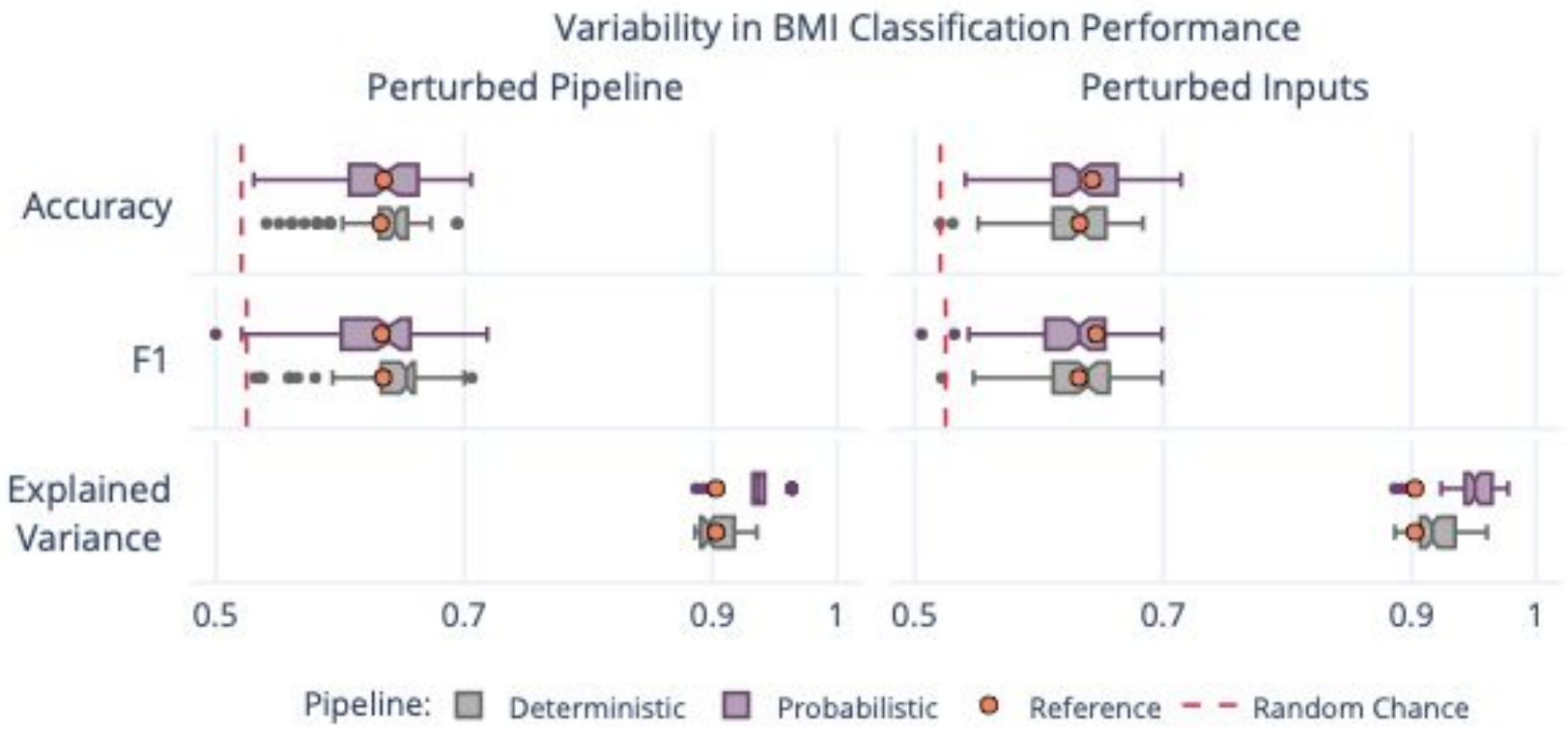
… and reducing differences across similar data (H_2, H_3)

(Kiar, 2020)

Chapter 4: Augmenting Connectomics Datasets Through Aggregation of Numerically Unstable Derivatives

In progress (to be submitted to IEEE Transactions on Medical Imaging)

Considerable Uncertainty in Classification



(Kiar, 2020; associated with previous chapter)

Outline of Experiment

- Using the existing MCA dataset...
- And 3 different measures of high/low fitness...
  (BMI, VO_2Max, Cholesterol)
- Train 3 classifiers with various forms of aggregation*...
  (SVM, RF, LRC)
- Evaluate classification performance w.r.t aggregation.

https://docs.google.com/document/d/13oazYfnbFRGC2fT-PfjQbYIq0BU2atzFXWGs2tfj7ZY/edit?usp=sharing

Aggregation Methods

A1.	(No aggregation)	jackknife resampling
A2.	(Naive)	edgewise mean of connectomes
A3.	(Naive)	edgewise median of connectomes
A4.	(Data Aware)	distance-dependent consensus
A5.	(Meta-Analysis)	vote across jackknifed classifiers
A6.	(Mega-Analysis)	group and train on all connectomes
B1.	(Multi-Acquisition)	A{0-5} across same-subject
B2.	(Multi-Subsample)	A{0-5} across same-scan

https://docs.google.com/document/d/13oazYfnbFRGC2fT-PfjQbYIg0BU2atzFXWGs2tfj7ZY/edit?usp=sharing

Other Publications

List of Co-Authored Publications

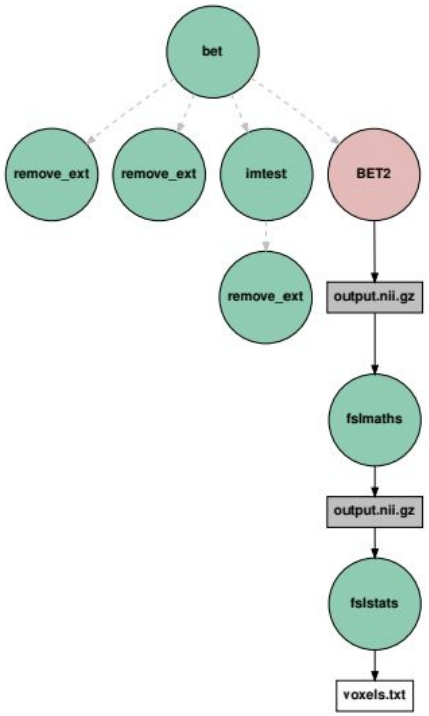
1. Membrane Protein Classification: a Reproducibility Study. H. Heidarzadeh, G. Kiar, T. Glatard. Work in progress (2020).
2. Identifying the Source of Instabilities Within Neuroimaging Pipelines. Y. Chatelain, G. Kiar, A. Salari, T. Glatard. Work in progress (2020).
3. Exploring the Relationship Between Early Psychosis Verbal Memory Deficits and White Matter Integrity. C. Henri-Bellemare, ..., G. Kiar, ..., M. Lepage. Submitted, Schiz. (2020).
4. File-based localization of numerical perturbations in data analysis pipelines. A. Salari, G. Kiar, ..., T. Glatard. GigaScience. (2020).
5. Neural correlates of polygenic risk score for autism spectrum disorders in the general population. B. Khundrakpam, ..., G. Kiar, ..., A. C. Evans. Brain Communications (2020).
6. qEEG toolbox for the MNI Neuroinformatics ecosystem: normative SPM of EEG source spectra. J. Bosch-Bayard, ..., G. Kiar, ..., P. Valdes-Sosa. Front. in Neuroinf. (2020).
7. Deploying Large Fixed File Datasets with SquashFS and Singularity. P. Rioux, G. Kiar, ..., S. T. Brown. PEARC '20, Association for Computing Machinery. (2020).
8. Brain status modeling with non-negative projective dictionary learning. M. Zhang, ..., G. Kiar, ..., A. C. Evans. Neuroimage. (Oct. 2019).
9. PyBIDS: Python tools for BIDS datasets. T. Yarkoni, ..., G. Kiar, ..., R. Blair. JOSS. (2019).
10. Boutiques: a flexible framework to integrate command-line applications in computing platforms. T. Glatard, G. Kiar, ..., A. C. Evans. GigaScience 7.5. (2018).

List of Co-Authored Publications

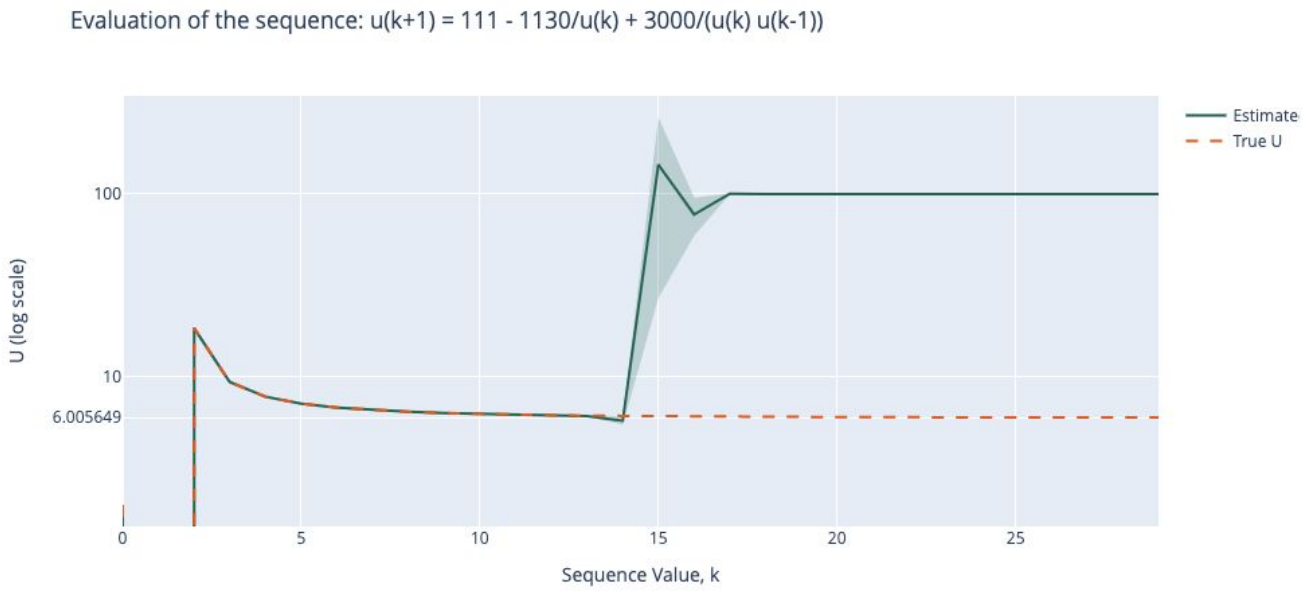
1. Membrane Protein Classification: a Reproducibility Study. H. Heidarzadeh, G. Kiar, T. Glatard. Work in progress (2020).
2. Identifying the Source of Instabilities Within Neuroimaging Pipelines. Y. Chatelain, G. Kiar, A. Salari, T. Glatard. Work in progress (2020).
3. Exploring the Relationship Between Early Psychosis Verbal Memory Deficits and White Matter Integrity. C. Henri-Bellemare, ..., G. Kiar, ..., M. Lepage. Submitted, Schiz. (2020).
4. File-based localization of numerical perturbations in data analysis pipelines. A. Salari, G. Kiar, ..., T. Glatard. GigaScience. (2020).
5. Neural correlates of polygenic risk score for autism spectrum disorders in the general population. B. Khundrakpam, ..., G. Kiar, ..., A. C. Evans. Brain Communications (2020).
6. qEEG toolbox for the MNI Neuroinformatics ecosystem: normative SPM of EEG source spectra. J. Bosch-Bayard, ..., G. Kiar, ..., P. Valdes-Sosa. Front. in Neuroinf. (2020).
7. Deploying Large Fixed File Datasets with SquashFS and Singularity. P. Rioux, G. Kiar, ..., S. T. Brown. PEARC '20, Association for Computing Machinery. (2020).
8. Brain status modeling with non-negative projective dictionary learning. M. Zhang, ..., G. Kiar, ..., A. C. Evans. Neuroimage. (Oct. 2019).
9. PyBIDS: Python tools for BIDS datasets. T. Yarkoni, ..., G. Kiar, ..., R. Blair. JOSS. (2019).
10. Boutiques: a flexible framework to integrate command-line applications in computing platforms. T. Glatard, G. Kiar, ..., A. C. Evans. GigaScience 7.5. (2018).

Spotlight: Finding Instabilities in Pipelines

Detecting OS Differences



Finding Unstable Operations

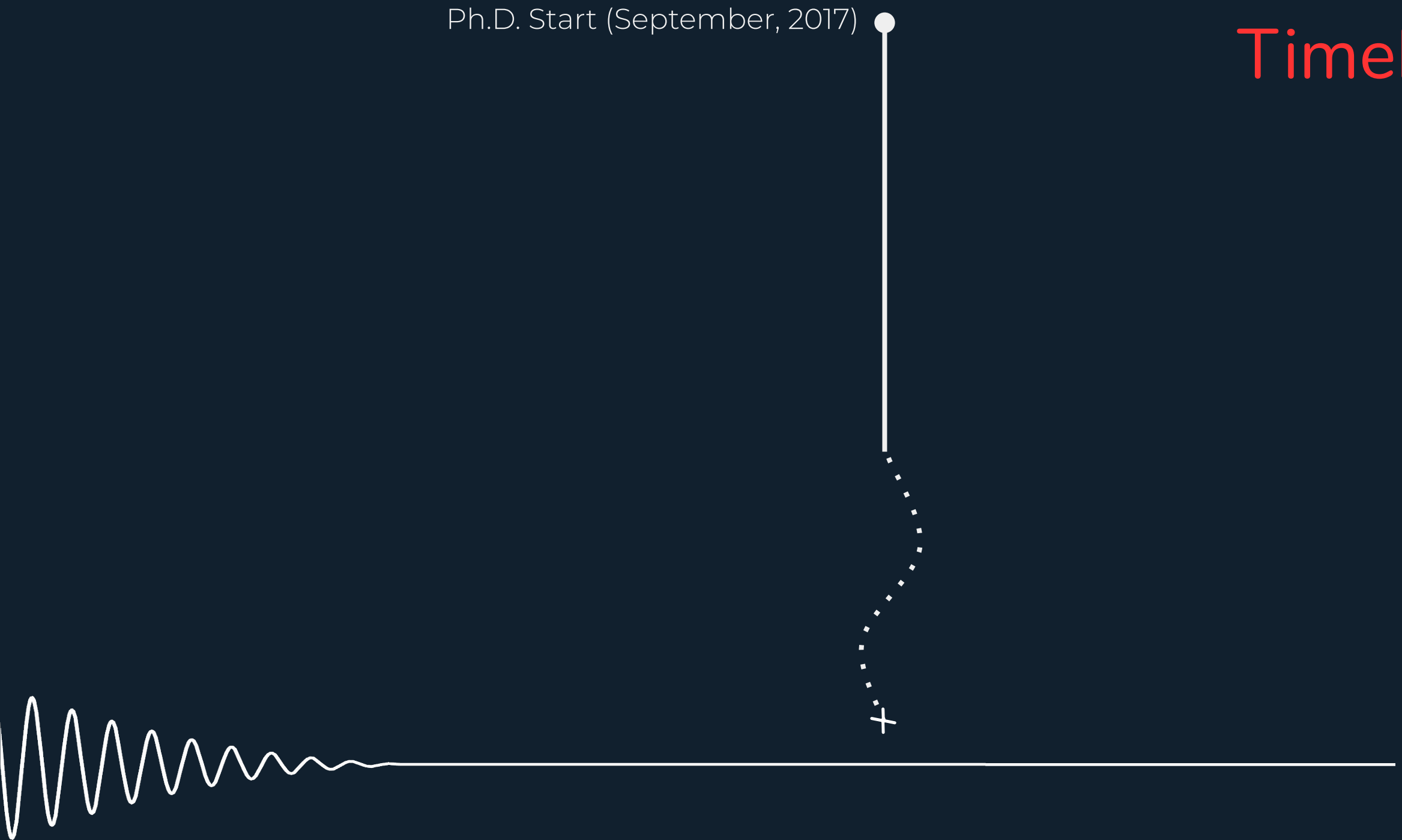


(Salari, 2020)

Collaboration with Yohan Chatelain, Ph.D.

Timeline

Timeline



Timeline

Ph.D. Start (September, 2017)

NSERC CGS-D (March, 2018)

Boutiques (2nd author; March, 2018)

Year 0 – 1

Timeline

Ph.D. Start (September, 2017)

NSERC CGS-D (March, 2018)

Boutiques (2nd author; March, 2018)

Clowdr (March, 2019)

Year 0 – 1

Year 1 – 2

Timeline

Ph.D. Start (September, 2017)

NSERC CGS-D (March, 2018)

Boutiques (2nd author; March, 2018)

Year 0 – 1

Clowdr (March, 2019)

Year 1 – 2

Comparing Perturbations (November, 2019)

Comparing Perturbations (March, 2020)

CONP Research Fellowship (July, 2020)

Analytic Impact of Instability (September, 2020)

Year 2 – 3

Timeline

Ph.D. Start (September, 2017)

NSERC CGS-D (March, 2018)
Boutiques (2nd author; March, 2018)

Year 0 – 1

Clowdr (March, 2019)

Year 1 – 2

Comparing Perturbations (November, 2019)

Comparing Perturbations (March, 2020)

CONP Research Fellowship (July, 2020)

Analytic Impact of Instability (September, 2020)

Year 2 – 3

Aggregating Derivatives (November, 2020)

Analytic Impact of Instability (December, 2020)

Write Thesis (December, 2020)

Submit Thesis (February, 2021)

Year 3 – 3.5

Timeline

Ph.D. Start (September, 2017)

NSERC CGS-D (March, 2018)

Boutiques (2nd author; March, 2018)

Clowdr (March, 2019)

Comparing Perturbations (November, 2019)

Comparing Perturbations (March, 2020)

CONP Research Fellowship (July, 2020)

Analytic Impact of Instability (September, 2020)

Aggregating Derivatives (November, 2020)

Analytic Impact of Instability (December, 2020)

Write Thesis (December, 2020)

Submit Thesis (February, 2021)

Ph.D. Defence (March, 2021)

Year 0 – 1

Year 1 – 2

Year

Year 3 – 3.

Total Time: 3.5 years

All code and data derivatives mentioned in this presentation are publicly available.

Thanks!

Find me @

gkiar

g_kiar

greg.kiar@mail.mcgill.ca

Acknowledgements

Fondation Brain Canada Foundation
HEALTHY BRAINS FOR HEALTHY LIVES
NSERC CRSNG
CANADA FIRST RESEARCH EXCELLENCE FUND
APOGÉE CANADA FONDS D'EXCELLENCE EN RECHERCHE
Mitacs Globalink
Inria inventeurs du monde numérique

...

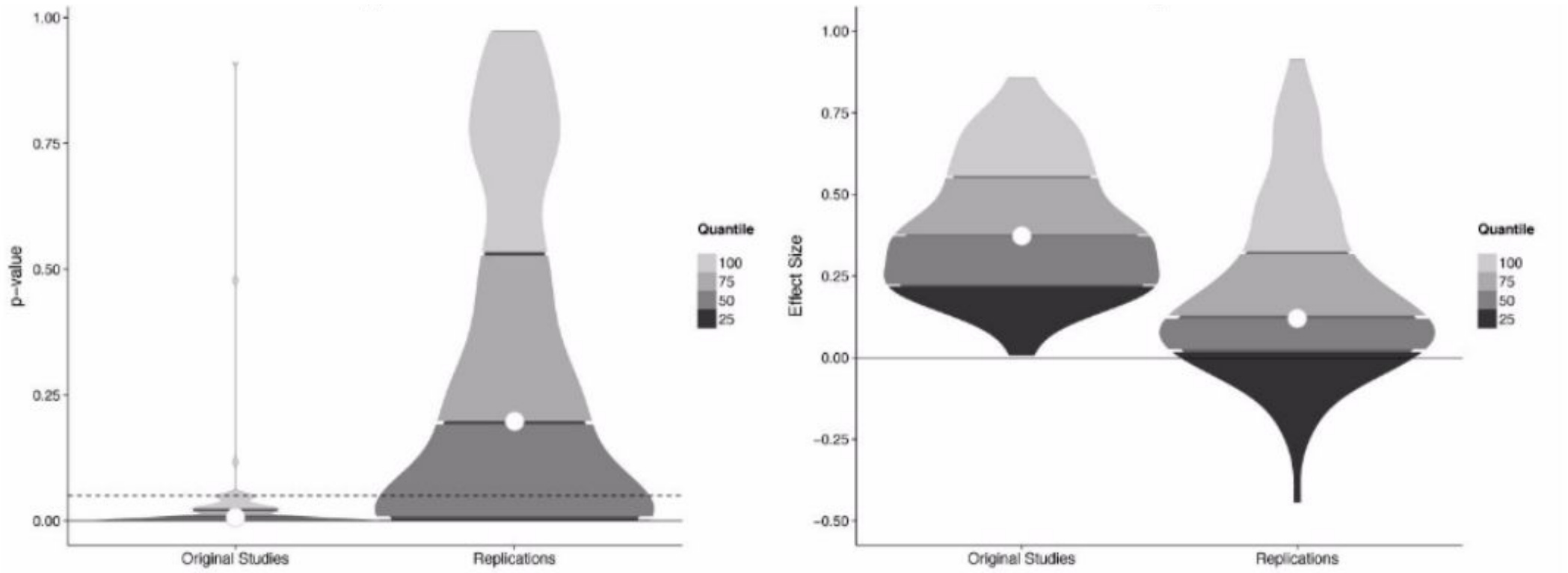
Questions?

(unorganized) Backup slides below

Issues of Reproducibility in Neuroscience

- Noisy data and incomplete statistics can lead to spurious results (Bennett et al., 2011) (fMRI)
- Operating system differences have led to different results (Glatard et al., 2015) (sMRI)
- Dominant software libraries have inflated false-positive rates (Eklund et al., 2016) (fMRI)
- 1-voxel perturbations to inputs result in significantly different outputs (Lewis et al., 2016) (sMRI)
- Similar tools performing similar operations give different results (Bowring et al., 2019) (fMRI)

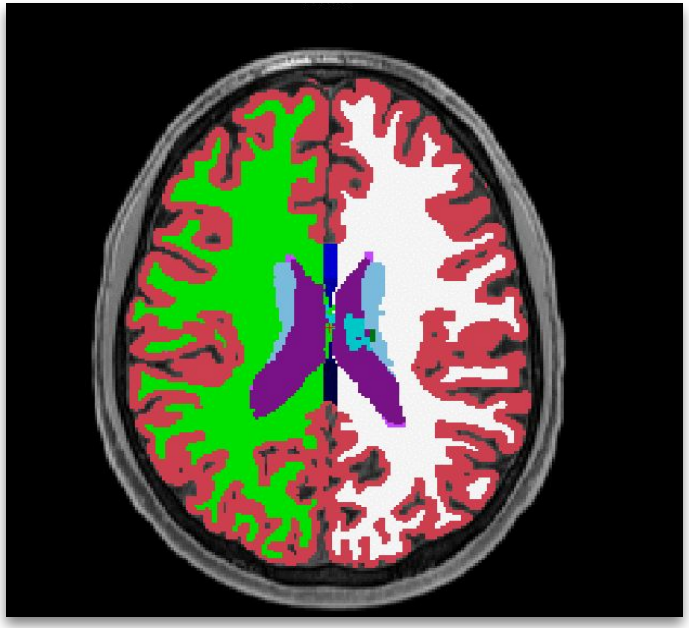
Replicability is Measurable

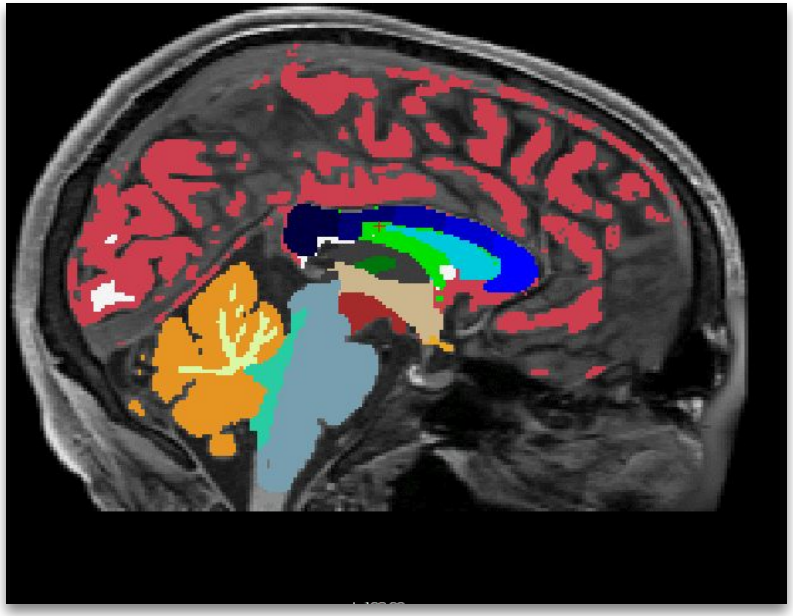


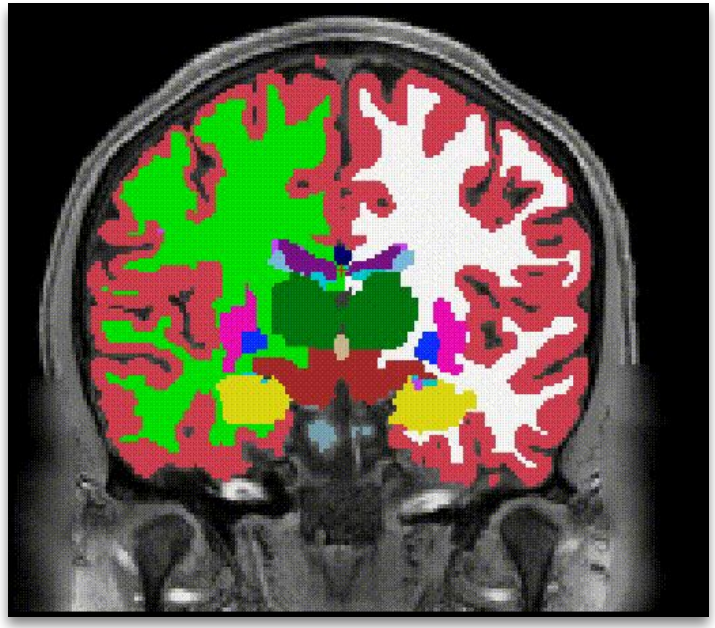
(Open Science Collaboration, 2015)

Difference Spotted! OS

CentOS 6 vs CentOS 7





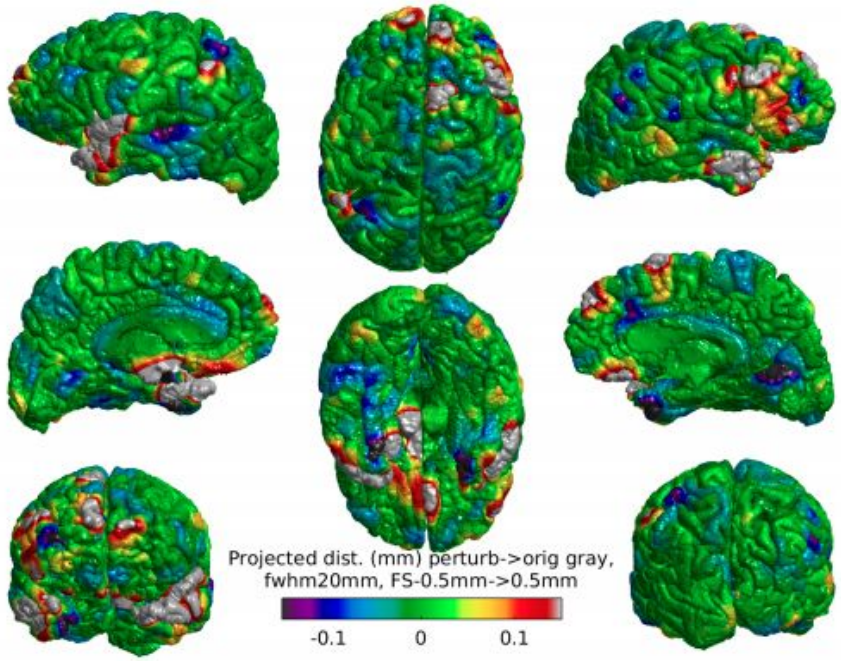


(Scaria, 2017)

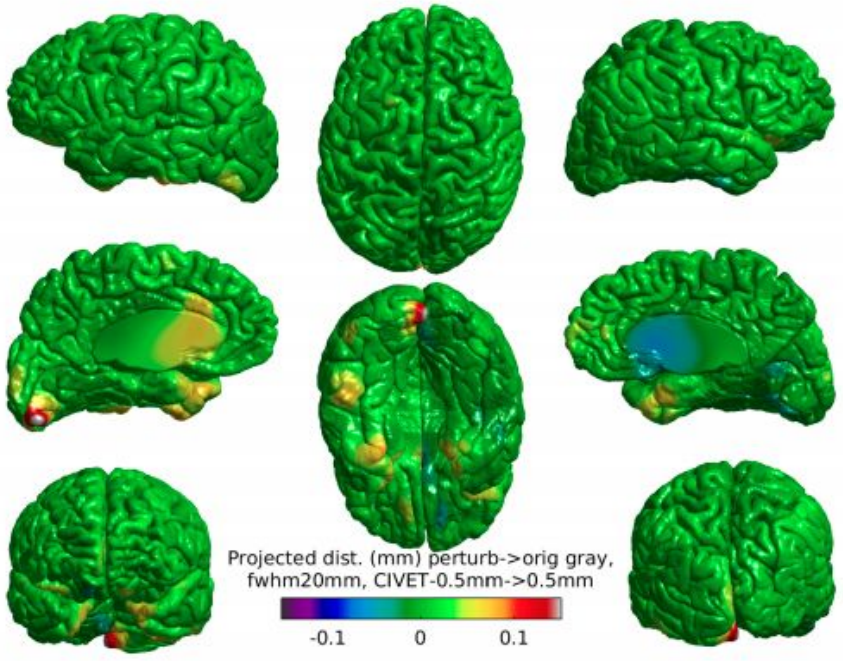
(Scaria, 2017)

Difference Spotted! Instability

1-voxel noise injections at 1% intensity



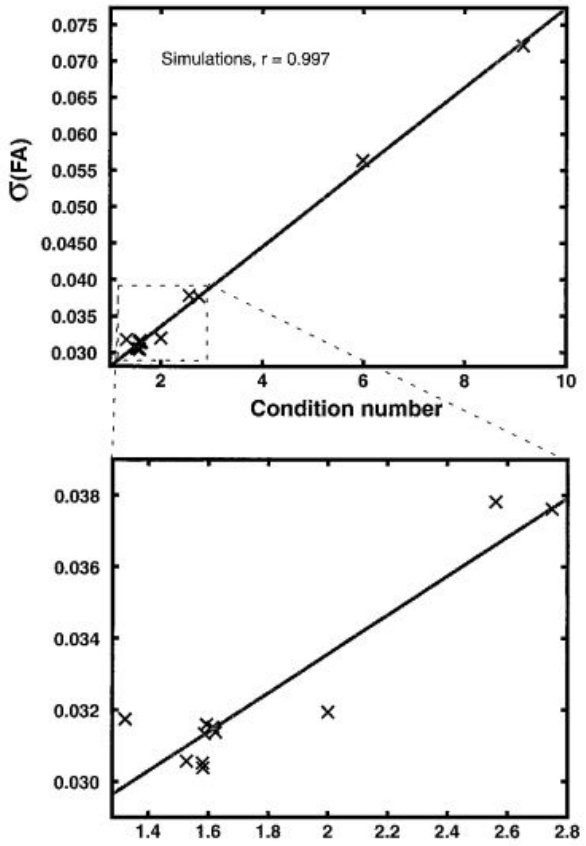
Freesurfer

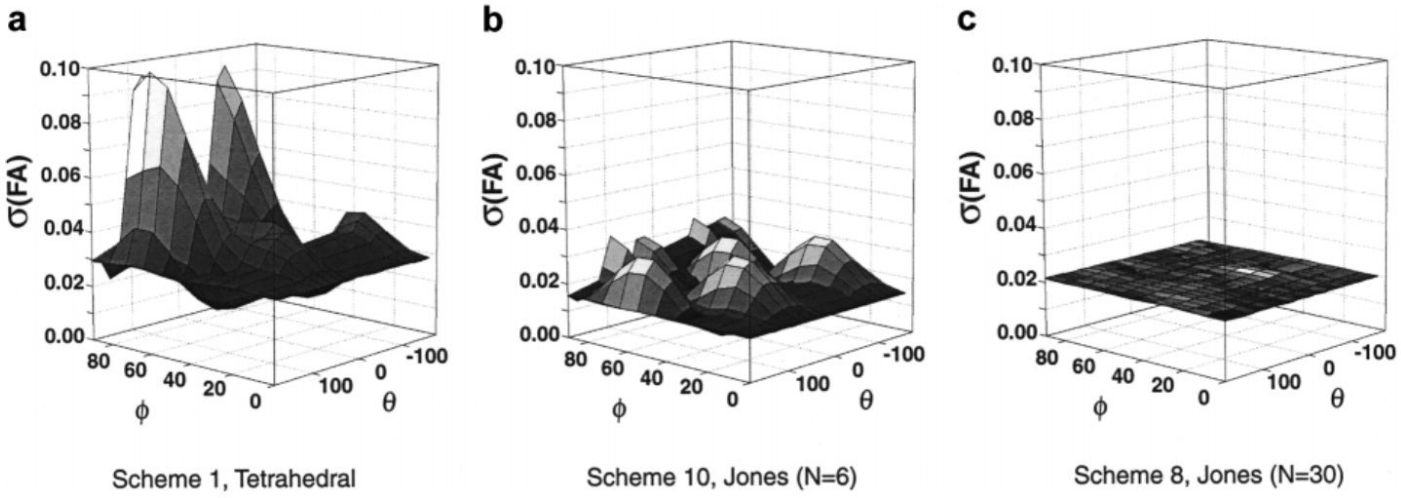


CIVET

(Lewis, 2017)

Instabilities can* be Anticipated



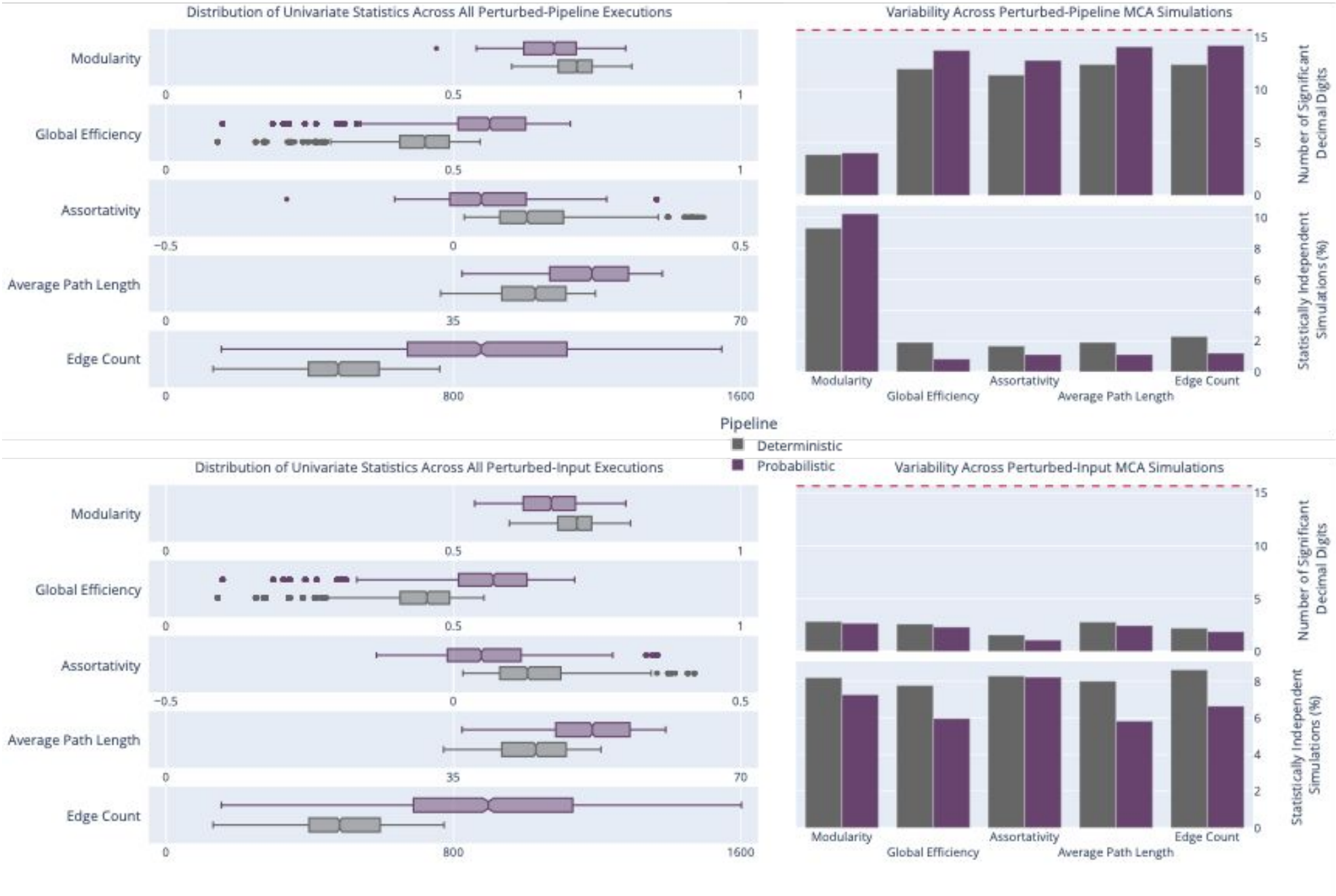


*sometimes

(Skare, 2000)

Discrim. Takeaway(s):

1) We can always tell data apart, even when that isn't desired.
2) Derivatives are highly discriminable -- there is a strong relationship between input and output data.
3) Input-type perturbations minimize acquisition variation.


Distribution of Univariate Statistics Across All Perturbed-Pipeline Executions
Modularity
Global Efficiency
Assortativity
Average Path Length
Edge Count
Variability Across Perturbed-Pipeline MCA Simulations
Number of Significant Decimal Digits
Statistically Independent Simulations (%)
Pipeline
Deterministic
Probabilistic
Distribution of Univariate Statistics Across All Perturbed-Input Executions
Variability Across Perturbed-Input MCA Simulations

Univariate Takeaway(s):

1) At a population level, both perturbation methods leave the distribution of statistics largely unchanged
2) At an individual level, input-type perturbations dramatically reduce the number of significant digits

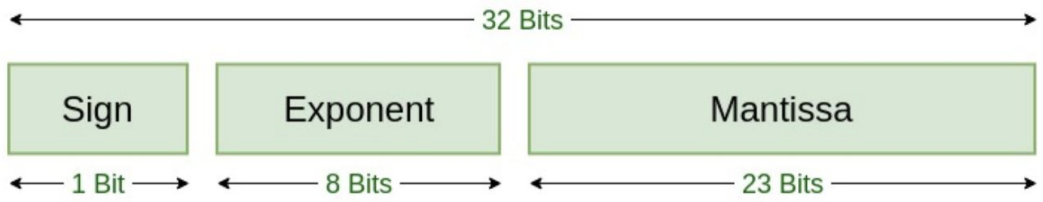
What I've learned in a sentence:

There is no "silver bullet statistic" or single measure we can trust to show us whether or not our experiments are stable, implying that we must evaluate everything, always.

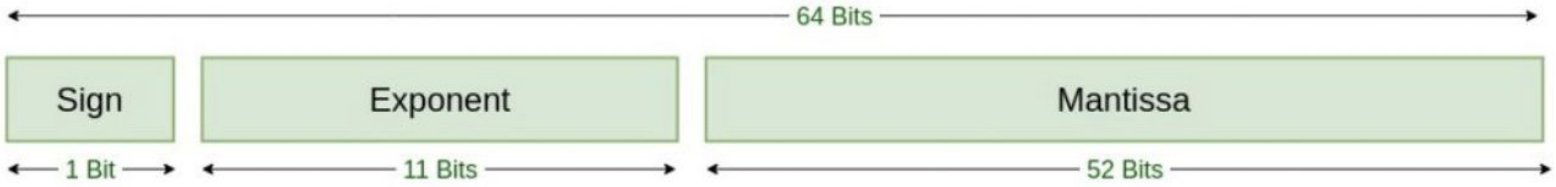
how MCA works

(with an example)

Floating Point Data are Finite



Single Precision
IEEE 754 Floating-Point Standard



Double Precision
IEEE 754 Floating-Point Standard

https://www.geeksforgeeks.org/ieee-standard-754-floating-point-numbers/

Floating Point Arithmetic is Inexact

E.g. addition is non associative

<u>Let's say we have 4 decimal digits of precision:</u>

$$(2000. \oplus -1998.) \oplus 1.333 = 2.000 \oplus 1.333 = 3.333$$

$$2000. \oplus (-1998. \oplus 1.333) = 2000. \oplus -1997. = 3.000$$

What is the right answer?

(inspired by Parker et al., 1997)

Floating Point Arithmetic is Inexact

E.g. addition is non associative

Catastrophic Cancellation

$$(2000. \oplus -1998.) \oplus 1.333 = 2.000 \oplus 1.333 = 3.333$$

$$2000. \oplus (-1998. \oplus 1.333) = 2000. \oplus -1997. = 3.000$$

(inspired by Parker et al., 1997)

Floating Point Arithmetic is Inexact

E.g. addition is non associative

$$(2000. \oplus -1998.) \oplus 1.333 = 2.000 \oplus 1.333 = 3.333$$

$$2000. \oplus (-1998. \oplus 1.333) = 2000. \oplus -1997. = 3.000$$

Round-off Error

(inspired by Parker et al., 1997)

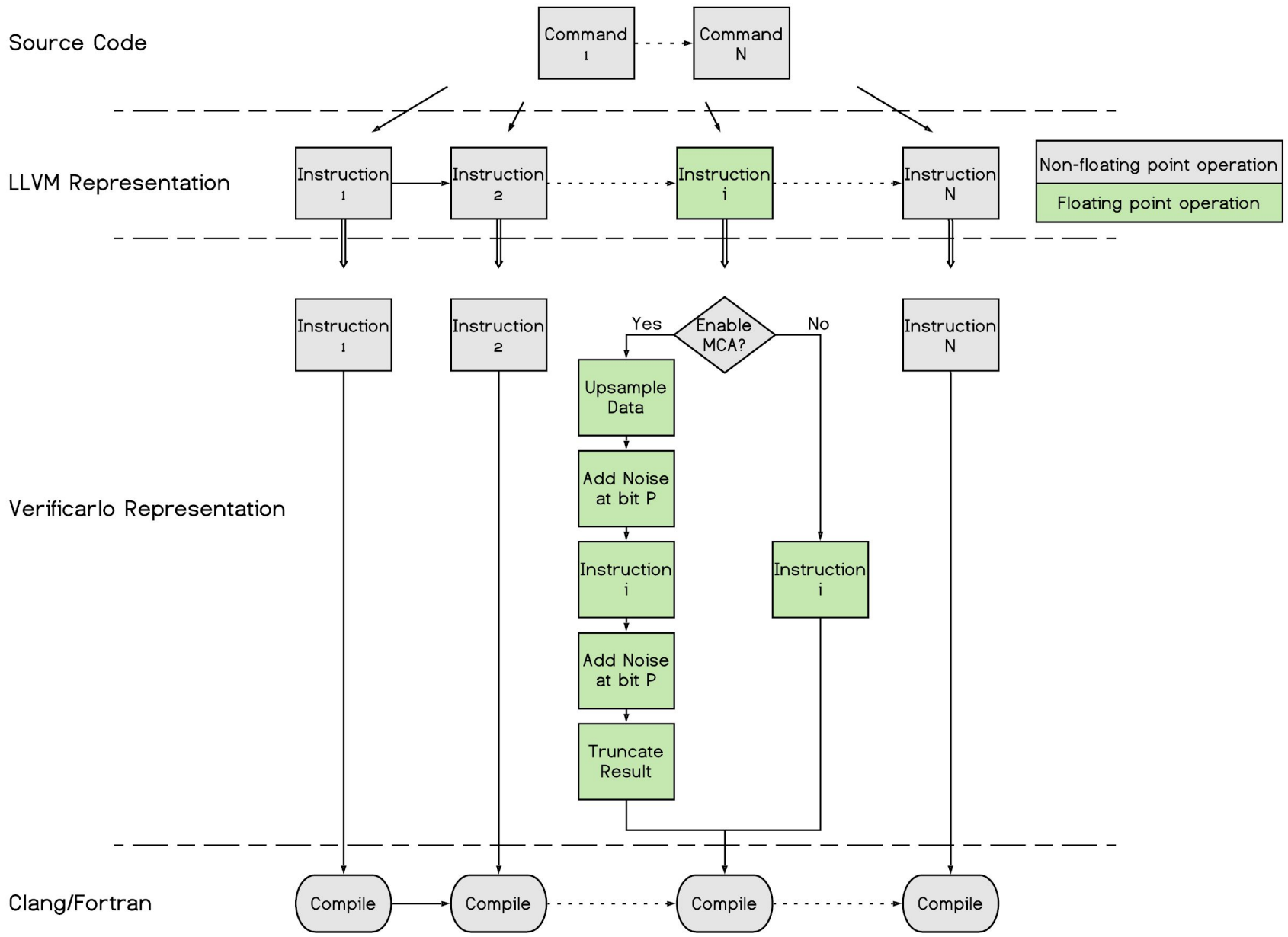
Monte Carlo Arithmetic (MCA)

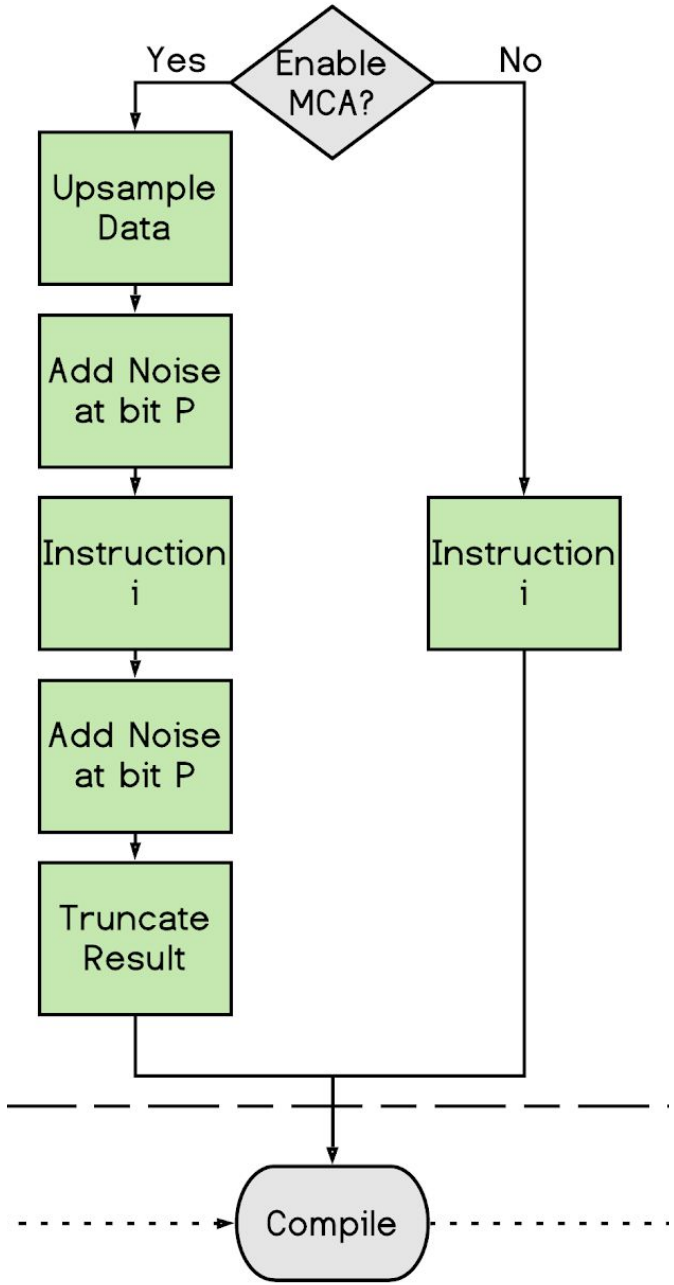
Inexact FP quantities become random variables

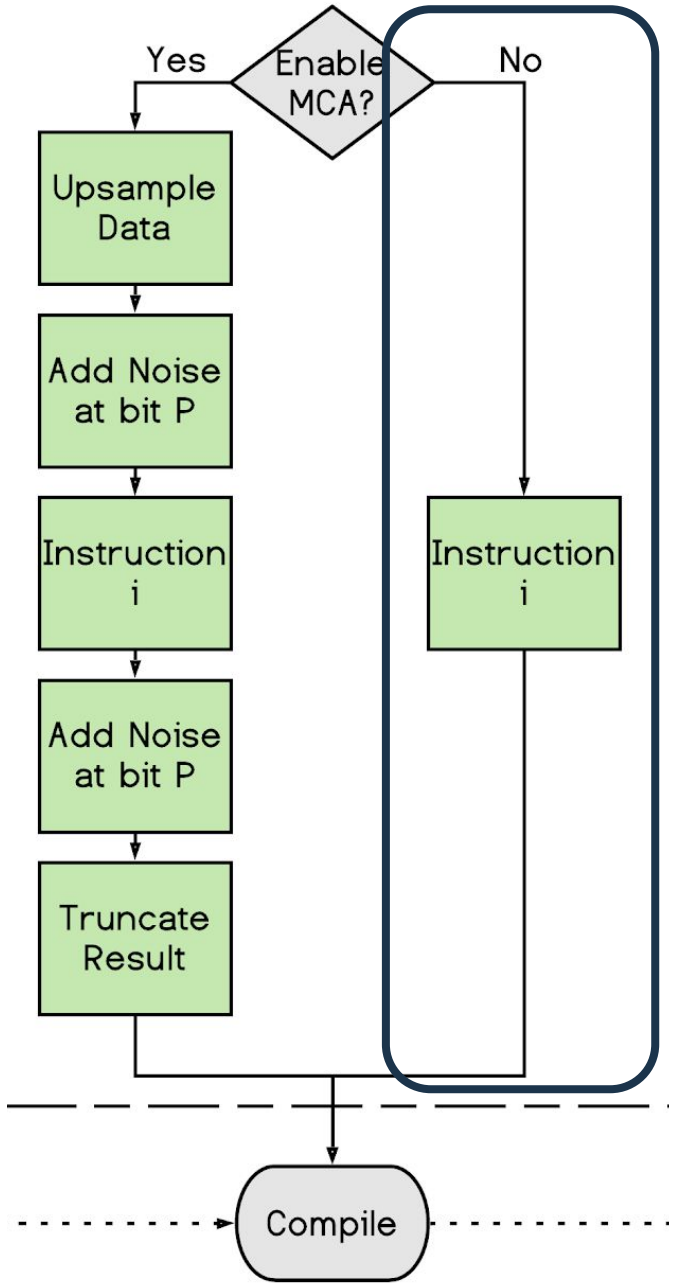
$$\tilde{x} \quad = \quad \textsf{inexact}\,(x,\, s,\, \xi) \quad = \quad x \;+\; 2^{e-s}\,\xi \qquad \text{where } e \text{ is the order of magnitude of } x$$

$$t\_\textsf{digit\_precision}\,(x) \quad = \quad \begin{cases} x & \text{if } x \text{ can be expressed exactly with } t \text{ digits} \\ \textsf{inexact}\,(x,\, t,\, \xi) & \text{otherwise.} \end{cases}$$

(Parker et al., 1997)


Source Code
Command 1
Command N
LLVM Representation
Instruction 1
Instruction 2
Instruction i
Instruction N
Non-floating point operation
Floating point operation
Verificarlo Representation
Instruction 1
Instruction 2
Yes
Enable MCA?
No
Instruction N
Upsample Data
Add Noise at bit P
Instruction i
Add Noise at bit P
Truncate Result
Instruction i
Clang/Fortran
Compile
Compile
Compile
Compile


Yes
Enable MCA?
No
Upsample Data
Add Noise at bit P
Instruction i
Add Noise at bit P
Truncate Result
Instruction i
Compile

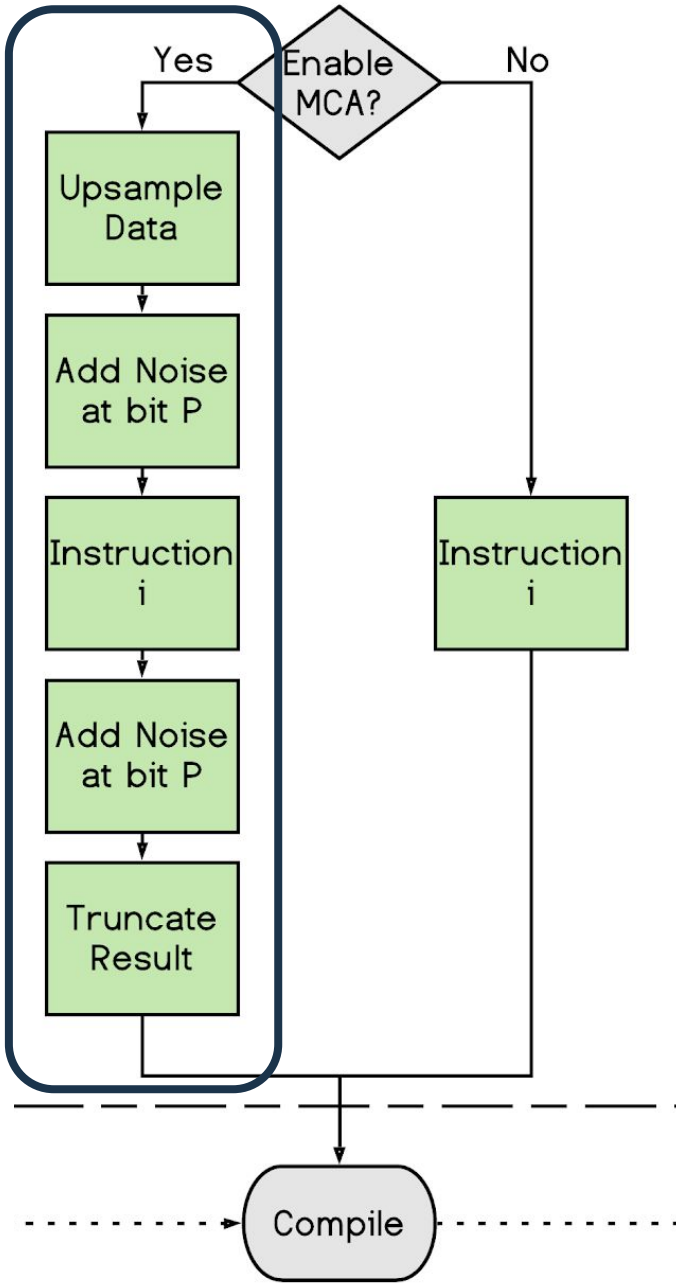


6000. + 5.452 = ...

6000. + 5.452

↳ 6005.

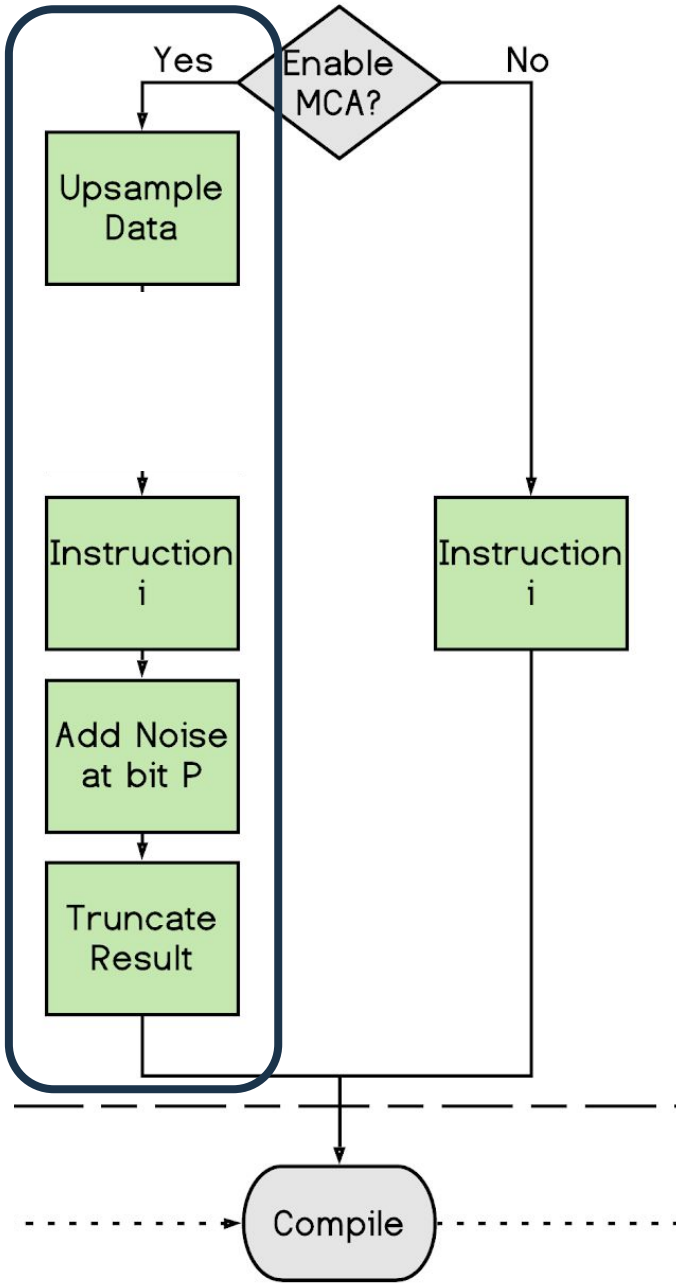
Native



6000. + 5.452 = ...

	6000.	+	5.452
Upsample:	↳ 6000.0000		↳ 5.4520000
	6000.0000		5.4520000
Perturb (PB):	↳ 6000.4293		↳ 5.4519512
	6000.4293	+	5.4519512
Operate:	↳6005.8813		
	6005.8813		
Perturb (RR):	↳ 6005.4924		
	6005.4924		
Truncate:	↳ 6005.		

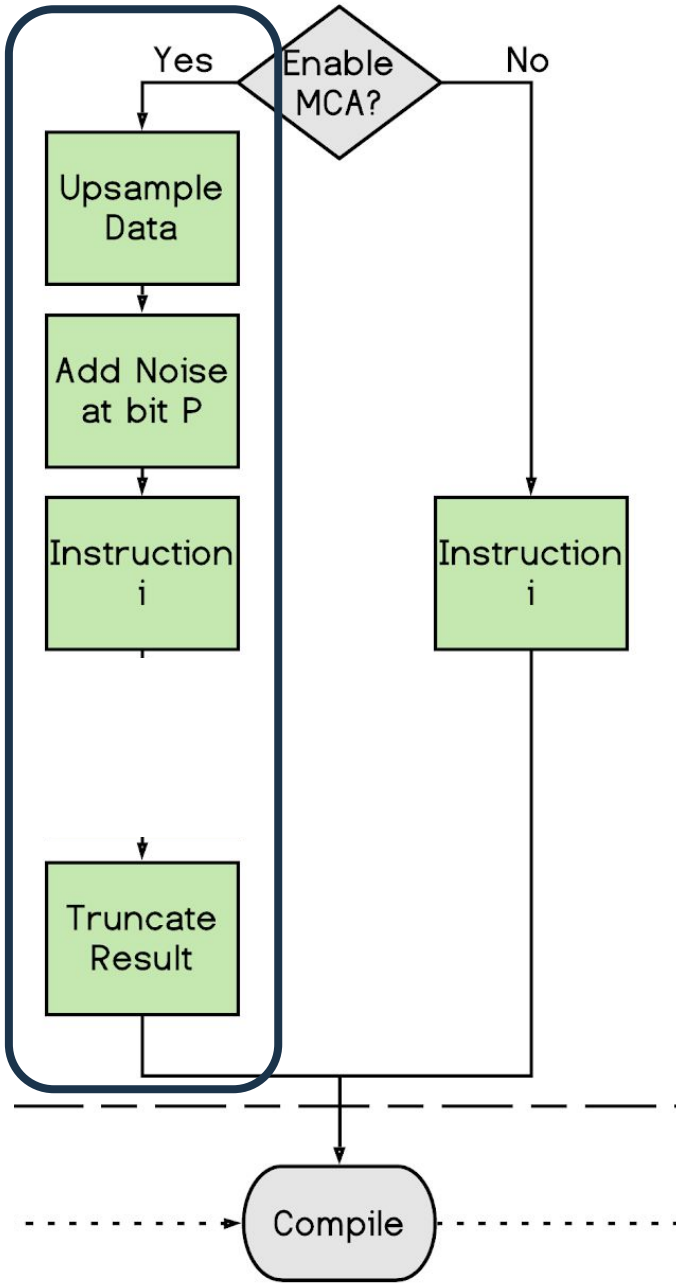
(full) MCA



6000. + 5.452 = ...

	6000.	+	5.452
Upsample:	↳ 6000.0000		↳ 5.4520000
	6000.0000	+	5.4520000
~~Perturb (PB):~~	~~↳ 6000.4293~~		~~↳ 5.4519512~~
	~~6000.4293~~	~~+~~	~~5.4519512~~
Operate:	↳6005.4520		
	6005.4520		
Perturb (RR):	↳ 6005.5274		
	6005.5274		
Truncate:	↳ 6006.		

RR



6000. + 5.452 = ...

	6000.	+	5.452
Upsample:	↳ 6000.0000		↳ 5.4520000
	6000.0000		5.4520000
Perturb (PB):	↳ 6000.4293		↳ 5.4519512
	6000.4293	+	5.4519512
Operate:	↳6005.8813		
	6005.8813		
~~Perturb (RR):~~	~~↳ 6005.4924~~		
	~~6005.4924~~		
Truncate:	↳ 6006.		

PB